AWARD NUMBER: W81XWH-16-1-0552

TITLE: Improving Diagnostics and Treatments for GWI Females by Accounting for the Effects of PTSD

PRINCIPAL INVESTIGATOR: Travis J.A. Craddock, Ph.D.

RECIPIENT: Nova Southeastern University
Fort Lauderdale, FL, 33314-7796

REPORT DATE: October 2017

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
October 2017	Annual	30 Sep 2016 - 29 Sep 2017

4. TITLE AND SUBTITLE
Improving Diagnostics and Treatments for GWI Females by Accounting for the Effects of PTSD

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-16-1-0552

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Craddock, TJA; Messer, S; Broderick, G; Klimas, NG

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Nova Southeastern University
3301 College Avenue
Fort Lauderdale, Florida, 33314-7796

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Material Command
Fort Detrick, Maryland, 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Our objective to disentangle the effects of post-traumatic stress disorder (PTSD) from Gulf War Illness (GWI) to understand the role that post-traumatic stress plays in the symptoms and perpetuation of GWI in female veterans. This is being accomplished by performing a secondary systems biology analysis for female subjects based on resources, methodology and results from our ongoing GWI research initiatives.

We have made substantive progress towards Year 1 objectives including consolidation and evaluation of trauma scores in GWI symptoms measures, subtyping of GWI subjects based on a Davidson Trauma Score cutoff, and preliminary statistical analysis of symptom measures in subtypes compared to healthy controls supporting the hypothesis that both subgroups are distinct from controls.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
					19b. TELEPHONE NUMBER *(Include area code)*

Reset

TABLE OF CONTENTS

INTRODUCTION:

As there is no widely accepted biomarker for Gulf War Illness (GWI) afflicted veterans are commonly diagnosed based on psychological or psychiatric evaluation (Nicolson, 2001). Confounding and/or exacerbating factors often complicate this diagnosis. Veterans suffering from GWI are often also diagnosed with post-traumatic stress disorder (PTSD) (Nicolson, 2000). Indeed a recent population-based study indicates that 35% of veterans with GWI are estimated to also suffer from PTSD (Weiner et al., 2011). The overall effects of a PTSD diagnosis on the clinical presentation and potential treatment of GWI have not been evaluated. PTSD has also been associated with physical health comorbidities including greater musculoskeletal pain, and cardio-respiratory complaints (Pacella et al., 2013). As GWI is a complex multisymptom illness characterized by autonomic dysfunction, musculoskeletal pain, cognitive dysfunction, chemical sensitivity, unexplained fatigue, sleep disturbance, and skin rash (Kang et al., 2003; Bourdette et al., 2001; Unwin et al., 1999), these studies indicate a strong overlap of PTSD with symptoms defining GWI, which complicates diagnosis and may affect potential treatment avenues. More specifically, subjects with PTSD exhibit immunological dysfunction, including increased circulating cytokines, lower natural killer cell activity, and lower total T lymphocyte counts when compared to healthy controls. As GWI has also been shown to have similar immunological underpinnings this suggests that a screen for probable PTSD diagnosis may delineate at least two phenotypes of GWI. However, while data over the last decade has revealed that immunological dysfunction likely plays a key role in GWI, most evaluations of GWI have focused on males. Our recent research has revealed marked sex differences in the inflammatory dysfunction displayed by GWI veterans. Furthermore, examination of differential patterns of gene expression in immune cell populations in subjects with PTSD indicates a gender dimorphism in immune pathways activated in PTSD, further suggesting that immune regulation in GWI women is differentially modified in the presence of probable PTSD diagnosis compared to men. In this work we are building on our ongoing research directed at mapping complex inflammatory mechanisms in GWI to improve our understanding of the immunologic underpinnings of GWI in women, and the compounding effects of co-morbidity with PTSD. Furthermore, we will take advantage of our ongoing work in predictive modeling to assess possible changes to putative treatments of GWI in women in the context of probable PTSD diagnosis.

1. KEYWORDS:

Gulf War Illness (GWI); Post-traumatic Stress Disorder (PTSD); Bio-behavioral Profile; Endocrine-Immune Regulation; Systems Modeling; Treatment Prediction

2. ACCOMPLISHMENTS:

What were the major goals of the project?

The efforts during the first year of this project were focused on the following specific aims and major tasks.

Specific Aim 1: Review of cases to subtype groups and refine definition of GWI with PTSD

- Major Task 1 – Consolidate Data and Review Cases Assessments
  Estimated completion: 65%
- Major Task 2 - Significant changes in psychological profiles.
  Estimated completion: 40%

Specific AIM 2. ***Identify and validate bio-behavioral patterns that confirm diagnosis.***

- *Major Task 1 - Significant changes in biological profiles.*
  Estimated completion: 10%

Specific AIM 3. Simulate optimal subtype specific treatments

- *Major Task 1 – Refine Endocrine-Immune Wiring Model.*
  Estimated completion: 100%
- Major Task 3 – Identify drug targets
  Estimated completion: 20%

What was accomplished under these goals?

Recruitment:

- Recruited **Mary Jeffrey** (M.A. in Clinical Psychology, 2015, University of Colorado), effective October 24, 2016, into position of Data Analyst I. Supported under the concurrent sister project W81XWH-16-1-0632 for 50%.
- Recruited **Rajeev Jaundoo** (B.Sc. in Psychology, 2015, Nova Southeastern University), effective November 9, 2016, into position of Research Programmer I. Supported under the concurrent sister project W81XWH-16-1-0632 for 37%.
- Recruited **Ricardo Castellanos** (M.Sc. in Electrical and Computer Engineering, 2015, Florida Atlantic University), effective December 12, 2016, into position of Research Associate I. Supported under the concurrent sister project W81XWH-16-1-0632 for 50%.
- Recruited **Ramesh Govindan** (B.A. in Biomedical Engineering, 2013, Dartmouth College), effective December 14, 2016, into position of Data Analyst I. Supported under the concurrent sister project W81XWH-16-1-0632 for 50%.
- Recruited NSU Psy.D. student **Rebecca Reinhardt** (M.S. in Forensic Psychology, American International College, 2015) effective January 30, 2017 into position of Research Assistant.

- Recruited NSU Psy.D. student **Jessica McDonald** (B.S., Psychology & B.A. Sociology, 2013; Military Social Work Graduate Certificate, 2015, University of Central Florida) effective January 30, 2017 into position of Research Assistant.

Specific Aim 1: Review of cases to subtype groups and refine definition of GWI with PTSD.

Major Task 1 – Consolidate Data and Review Cases Assessments

Subtask 1: Consolidate Data

Under the primary direction of Dr. Messer, and oversight by Dr. Craddock, Data Analyst Mary Jeffrey, with aid from Research Assistants Jessica McDonald and Rebecca Reinhardt, performed the following activities for the completion of Specific Aim 1 - Major Task 1 – Subtask 1:

- Primary study data were identified from the NSU-INIM archives.
- Hard-copies of the Davidson Trauma Scale (DTS), the self-administered Posttraumatic Stress Disorder (PTSD) diagnostic screening questionnaire, were checked for physical integrity and legibility. Less than a half-dozen of the packets required re-scanning to improve clarity of responses and/or copying.
- In addition to the DTS, hard-copies of the Composite International Diagnostic Interview (CIDI) computerized response forms were located, checked, and sorted.
- The DTS and CIDI forms were sorted by patient ID number in preparation for data entry.
- A secure Excel database with pre-populated variable entry fields and data range restrictions was constructed. Additionally, DTS and CIDI data element dictionaries were developed.
- Each DTS response field was entered into the database (i.e., both the identifying information fields [sex, age, etc.] and the symptom item frequency and severity responses). For the CIDI, similar respondent identifying information field data were entered as were responses to the PTSD symptom and diagnosis-related items (e.g., age of onset, duration) in Section K.
- A data quality check was conducted using manual double-key entry for all fields. The proportion of errors identified was less than 2 per 1000 data fields.
- The database and corresponding reference materials necessary was created to facilitate data entry amongst research staff. The database contains all variables from the Composite International Diagnostic Interview (CIDI) and the Davidson Trauma Scale (DTS), comprising approximately 1500 variables for each subject. The database incorporates cases from two separate studies (VA Merit (Klimas PI and W81XWH-09-2-0071 (Klimas PI)). The database is password protected and populated with de-identified data for safety, ethical and compliance purposes.
- The database currently houses 69 female subjects (19 GWI (GWI); 50 Healthy Controls (HC)).
- Consolidated entered data with past data collections efforts containing additional psychological and symptom based measures including the 36 Item Short Form Health Survey (SF36), Multidimensional Fatigue Inventory (MFI), Pittsburgh

Sleep Quality Index (PSQI), Sickness Impact Profile (SIP), Paced Auditory Serial Addition Test (PASAT), and Krupp Fatigue Severity Scale (FSS). Data entry was verified case-by-case with previously entered data.

Subtask 2: Perform re-assessments of CIDI and DTS scores, apply cut-offs and sub-type subjects

Under the primary direction of Dr. Messer, and oversight by Dr. Craddock, Data Analyst Mary Jeffrey, with aid from Research Assistants Jessica McDonald and Rebecca Reinhardt, performed the following activities for the completion of Specific Aim 1 - Major Task 1 – Subtask 2:

- DTS and CIDI data were identified from our archives and were thoroughly screened, entered, and data quality-checked. As noted below, we are extracting additional DTS data from other NSU-INIM primary studies to ensure that our proposed targeted sample sizes are attained.
- Determination of the CIDI primary variable, PTSD diagnostic status, and examination of the variable's properties in relation to the DTS probable diagnostic variable has been delayed. The ascertainment of CIDI diagnoses requires computerized scoring algorithms, referred to as the "CIDI-AUTO" program. CIDI-AUTO has not been commercially available for some time, though it has an established record of high quality performance and is still used for DSM-IV coding. Our team located a copy and we are in the process of validating its functioning (see below for additional information).
- We have successfully scored and examined the DTS variables assessing PTSD clinical status. We have computed DTS scores for item (i.e., symptom) frequency, severity, and total PTSD symptomatology. Additionally, scores for the DSM-IV PTSD re-experiencing, avoidance, and hyperarousal symptom clusters were calculated, as well as the Total DTS composite score, following standard scoring criteria (e.g., DTS Manual (1996)) and with the aid of SPSS version 24 statistical analysis software.
- DTS scores were computed and assessed across study groups (i.e., Condition: Healthy Controls (HC) vs. Gulf War Illness (GWI)).
- Frequency distributions and descriptive statistics were calculated for the DTS Total Score to evaluate the variable's properties in relation to the existing research literature examining optimal DTS cut-scores for the determination of probable PTSD diagnosis.
- Following the work of McDonald et al. (2014), we applied a DTS cut-score of 70 to our data to examine its utility in creating our target study groups (HC with low prior trauma (HC), GWI with low prior trauma (GWI_LT), & GWI with high prior trauma (GWI_HT)).
- For Females, using this cut-score resulted in the following cell sizes (*n*) and DTS *M* (*SD*): (1) HC with low trauma (*n* = 46, *M* = 5.13, *SD* = 11.38), (2) GWI_LT (*n* = 8, *M* = 32.13, *SD* = 27.00), and (3) GWI_HT (*n* = 11, *M* = 96.36, *SD* = 11.10). This is summarized in **Table 1**.

Table 1: Subtyping of GWI and HC subjects based on probable PTSD diagnosis

Condition	Number of Subjects
GWI	**19**
High Trauma (HT)	**11**
Military	2
Non-military	2
Unidentified	15
Low Trauma (LT)	**8**
HC	**49**
High Trauma (HT)	**3**
Military	0
Non-military	0
Unidentified	3
Low Trauma (LT)	**46**

Major Task 2 - Significant changes in psychological profiles.

Subtask 1: Univariate Analysis

Under the primary direction of Dr. Messer, and oversight by Dr. Craddock, Data Analyst Mary Jeffrey, performed the following activities for the completion of Specific Aim 1 - Major Task 2 – Subtask 1:

- Preliminary t-test statistics for group differences in psychological and symptom measures were run between the low trauma control group (HC), and the three classification of GWI: total GWI group (GWI_Tot), and the sub-typed high trauma (GWI_HT) and low trauma (GWI_LT) groups. Analyses were conducted to investigate general differences between the control group and GWI as well as between high and low trauma levels for female participants. In general splitting into subgroups increased the difference between GWI_HT and HC, while decreasing the difference between GWI_LT and HC compared to GWI_Tot (Figure 1). Despite this change significant differences were found for all measures between GWI total and subgroups and HC except for differences between GWI_LT and HC for SF36-Emotional Limits. Comparisons between GWI_HT and GWI_LT groups only showed differences in SF36-Emotional Limitations. These preliminary findings are inconclusive in regards to GWI_HT and GWI_LT in females reporting symptom differences indicating distinctions between these two subgroups, however both subgroups are distinct from HC. Continued analysis with increased numbers is expected to affect these results and will be completed in the second year of the project.

Specific AIM 2. Identify and validate bio-behavioral patterns that confirm diagnosis

Major Task 1 - Significant changes in biological profiles.

Consolidate and Quality Control biological data

Under the primary direction of Dr. Craddock Data Analyst I, Ramesh Govindan, performed the following activities towards the completion of Specific Aim 2 – Major Task 1:

- Compiled de-identified demographics information containing age, gender, race, and BMI, as well as symptom scores, received from the Klimas clinic via honest broker David Freeman. Data was compiled into a spreadsheet based on patient ID number series.
- FACS data from a subset of the VA Merit Award (Klimas PI) has been compiled and sorted, and is in the process of validation.
- ELISA cytokine measures from VA Merit Award (Klimas PI) have been compiled, sorted and is in the process of validation.
- As this project draws data from an on-going study (GW120045 (VA Merit: Klimas PI) for secondary analysis, we are awaiting for the recruitment, testing, and analysis of data from the parent award.

Subtask 1: Refine algorithms for univariate and multivariate analysis

Under the primary direction of Dr. Craddock the following tasks towards the completion of Specific Aim 2 – Major Task 1 – Subtask 1 were performed.

- Research Associate I, Ricardo Castellanos developed a Matlab based data importer to load data from the database to computational pipeline in a way that the user will be able to start the process with minimum interaction while being able to select specific measurements and/or times from the database. This simplifies processing, speeds up the pipeline system and reduces potential for human error. The system handles selection of data for cytokines, hormones psychological and symptom measures, and demographic data. A process to import gene expression data is currently being constructed.
- Research Associate I, Ricardo Castellanos has integrated a standard pipeline in Matlab to generate an automatic process to evaluate statistical analysis of the database that entails One-Sample T-Test, Two-Sample T-Test, Paired Test (specifically for the timed variables), Spearman Correlations taking into account multiple testing and controlling for false discovery rate (FDR), partial correlations with respective FDR as well.
- Research Programmer I, Rajeev Jaundoo automated the reading and preparation of gene expression data files. The process automatically merges Entrez gene identifiers with files containing microarray probe data identifiers and gene expression values, and outputs lines containing information on specific genes according to predefined modules.
- Volunteer Research Staff, Zeus Cortes wrote a script to read XML based network files from the KEGG Pathway database (Kanehisa et al., 2011) and convert them to our pipeline supported gene module file format. Module files currently

accessible now include gene sets based on the human protein-protein interaction network as defined by Suthram et al. (2010) and previously used in Craddock et al. (2015), the databases integrated in the Consensus Path Database (Kamburov et al., 2009; 2011; 2013), and the KEGG Pathway database (Kanehisa et al., 2011).

- Research Associate I, Ricardo Castellanos with aid from Volunteer Research Staff, Zeus Cortes created a script to import machine learning libraries from the program WEKA (Hall et al., 2009) into Matlab cross platform. WEKA is a collection of machine learning algorithms for data mining tasks with algorithms that can either be applied directly to a dataset or called from a Java code. This was explored for the purpose of merging into the computational pipeline for automatic classification of selected data, however it was determined that using Matlab's Classification Learner Application was more versatile and generalizable within the Matlab framework.

Subtask 3: Mutlivariate Analysis

Under the primary direction of Dr. Craddock, Research Associate I, Ricardo Castellanos with aid from Volunteer Research Staff, Zeus Cortes performed the following tasks in support of Specific Aim 2 – Major Task 1 – Subtask 1.

- Applied Unsupervised Machine Learning (UML) using WEKA (Hall et al., 2009) to identify symptom based subtypes to better understand the significant impact of symptom subtypes in GWI. Our preliminary analysis using the UML procedure of Expectation Maximization (EM)(Do & Batzoglou, 2008) of n = 27 male veterans with GWI, and n = 29 healthy male Gulf War era veteran controls (HC) has yielded two distinct GWI subgroups (log likelihood = -80.4 (lower is better)), and only a single group of HC at a comparable level of likelihood (log likelihood = -79.0). The two GWI subgroups differ in measures of the Davidson Trauma Score with one high trauma group and one low trauma group. The high group shows increased measures of fatigue, and sickness impact, and decreased emotional and physical well-being compared to the low trauma group.
- UML was followed with Supervised Machine Learning (SML) using WEKA to identify subtype specific bio-signatures based on cytokine and hormone data. Our preliminary analysis using the SML procedure of Decision Trees on these two subgroups of GWI show a minimum of 93% accuracy in classifying GWI subgroups versus HC, and identified abnormal levels of cortisol as the prime discriminator suggesting a role for HPA axis dysfunction. When the procedure was used on the GWI group as a whole the accuracy was reduced to ~64%, slightly better than chance, indicating improved discrimination when using subtypes. Male data was used in this analysis, however the computational framework is available for use on female data when it is supplied. In Year 2 of the project we aim to expand this analysis to the complete female dataset.

Specific AIM 3. Simulate optimal subtype specific treatments

Major Task 1 – Refine Endocrine-Immune Wiring Model

Subtask 1: Incorporate detailed immune wiring logic with extended HPA-HPG system

The following tasks were performed towards the completion of Specific Aim 3 – Major Task 1 – Subtask 1.

- In conjunction with work done under award W81XWH-15-1-0582 (Broderick PI/ Craddock PI) the Craddock lab constructed aggregate models of the regulatory biology captured so far in our basic coarse grain endocrine-immune models (Cradock et al., 2014) with the finer grain model of immune signaling (Fritsch et al., 2013) (Figures 2-5) for females. Four distinct models are created to capture the cycles and regulation of the female sex steroid system. The Broderick lab is refining modeling procedures to incorporate this information into one single comprehensive model.

Major Task 3 – Identify drug targets

Subtask 2: Cross reference identified gene sets against pharmacogenomics database

Under the primary direction of Dr. Craddock, Research Programmer I, Rajeev Jaundoo, performed the following tasks towards the completion of Specific Aim 3 – Major Task 3 – Subtask 2.

- Research Programmer I, Rajeev Jaundoo, amalgamated numerous drug-gene databases into one large database. The databases currently include the pharmacogenomics knowledgebase (PharmGKB), FDA's National Drug Code Directory (FDA NDC), DrugBank.ca, the toxin and toxin-target database (T3DB), Guide to Pharmacology (GtP), and the HUGO Gene Nomenclature Committee (HGNC). MATLAB scripts were created that allow future versions of each of the above datasets, and new databases to be easily appended to the consolidated database. After raw database information is downloaded, the amalgamation script parses the file to obtain information relevant to the dataset, and includes any combination of gene-drug-target interactions. This relevant information is then written to an output file and added to the overall dataset.

What opportunities for training and professional development has the project provided?

Nothing to Report.

How were the results disseminated to communities of interest?

Volunteer research assistant Zeus Cortes from the Miami-Dade College School of Engineering and Technology presented results of a summer project at the Miami Dade College Summer Research Institute's Science Technology Engineering and Math (STEM) Symposium. This project aimed at applying machine learning algorithms to detect underlying correlations in GWI and HC cytokine data at three time points collected over a graded exercise challenge. Male data was used in this analysis,

however the computational framework is available for use on female data when it is supplied.

1. *Zeus Cortes-Aguila, Ricardo Castellanos, Rajeev Jaundoo, Travis J. A. Craddock* ***Application of Machine Learning Techniques to Classify High/Low Trauma Gulf War Illness Patients Against Healthy Controls***, Miami Dade College Summer Research Institute's Science Technology Engineering and Math (STEM) Symposium, Tuesday, September 5th, 2017

Additionally, ongoing progress was shared with fellow researchers and with veteran advocates through the meeting schedule established for the Gulf War Illness Research Consortium. In addition as part of an outreach initiative to the broader research and patient community Drs. Broderick and Craddock delivered podium talks and participated in discussion panels during a one-day symposium held at Nova Southeastern University October 26, 2016, where the initiation and aims of this project were discussed.

What do you plan to do during the next reporting period to accomplish the goals?
During Year 2 of this award we will continue our consolidation of the female data from parent award GW120045 (VA Merit: Klimas PI) and supplement with data from the Boston Gulf War Illness Consortium (Sullivan PI) (See section 5), and continue the evaluation of the significant changes in psychological and biological profiles between sub-groups of GWI and HC through univariate and multivariate statistics, with the multivariate analysis being supplemented by machine learning classifications. In conjunction with the Broderick lab at Rochester General Hospital we will analyze regulatory behaviors of the refined endocrine-immune wiring model, and compare biological profiles of GWI sub-groups to model predictions. We will also compare gene expression profiles of GWI sub-groups to healthy controls and cross-reference differentially expressed gene modules with our expanded pharmacogenomics database to drug targets that are unique to GWI with and without past trauma.

4. IMPACT:

What was the impact on the development of the principal discipline(s) of the project?
Nothing to report.

What was the impact on other disciplines?
The automation of this computational pipeline is not disease-specific and is applicable for examination and studies of any complex illness. We are currently exploring opportunities to automate the reading and analysis of biological data to identify pathways and processes of interest, the generation of new regulatory models incorporating these pathways and processes, and the prediction of treatment targets and courses for application to other neuro-inflammatory illnesses such as Parkinson's

disease, Alzheimer's disease, or neuroinflammation resulting from traumatic brain injury.

What was the impact on technology transfer?
Nothing to report.

What was the impact on society beyond science and technology?
Nothing to report.

5. CHANGES/PROBLEMS:

Changes in approach and reasons for change
The tasks as originally outlined for Specific Aim 2 – Major Task 1 – Subtask 3 have not changed, however we have incorporated Machine Learning processes to supplement the originally proposed methods. Machine Learning was incorporated as it offers the advantage of automated, unbiased feature learning for classification based on symptoms, or biology.

Actual or anticipated problems or delays and actions or plans to resolve them
As this project draws data from an on-going study (GW120045 (VA Merit: Klimas PI) for secondary analysis, we are awaiting for the recruitment, testing, and analysis of data from the parent award. To control for any additional delays we have opened discussions with Dr. Kimberly Sullivan at Boston University for data sharing from the Boston Gulf War Illness Consortium (W81XWH-13-2-0072 (Sullivan PI)) to request demographic (gender, age, BMI, condition), PTSD probability (as determined by meeting the CAPS (Clinician-Administered PTSD Scale) criteria), symptom measures (SF36 – veterans, MFI, PSQI), and biological data at baseline (Cytokine, FACS, hormone) for GWI female subjects to meet our goal of 20 subjects per group.

Changes that had a significant impact on expenditures
In February 2017, Dr. Gordon Broderick accepted a position at Rochester General Hospital in Rochester, New York. Prior to departing Nova Southeastern University the Clinical Systems Biology Group (Drs. Broderick and Craddock) shared a pool of computational resources including MATALB licenses, and cycle time at the University of Miami's Center for Computational Science supercomputer Pegasus II. After Dr. Broderick's transition his resources were freed for use by the Craddock lab. The original award requested funds to expand upon these core resources, however due to the move of the Broderick lab and the release of the core resources, we are requesting a rebudget of these funds.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents
Nothing to Report

Significant changes in use or care of human subjects
Nothing to Report

Significant changes in use or care of vertebrate animals
Nothing to Report

Significant changes in use of biohazards and/or select agents
Nothing to Report

6. PRODUCTS:

Publications, conference papers, and presentations

- **Journal publications:** Nothing to report.

- **Other publications, conference papers and presentations**
  Zeus Cortes-Aguila, Ricardo Castellanos, Rajeev Jaundoo, Travis J. A. Craddock ***Application of Machine Learning Techniques to Classify High/Low Trauma Gulf War Illness Patients Against Healthy Controls***, Miami Dade College Summer Research Institute's Science Technology Engineering and Math (STEM) Symposium, Tuesday, September 5th, 2017
  (see Appendix C)

- **Other Products**
  Consolidated Gene-Drug Database
  -a manuscript describing the database is in preparation.

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Name:	Travis Craddock
Project Role:	PI
Research Identifier (e.g. ORCID ID):	
Nearest person month worked:	3
Contribution to Project	Overall direction of the project and coordination of efforts lead by Drs. Messer, Broderick and Klimas. Primary responsibility and leadership of pattern identification and treatment simulation.
Funding Support:	

Name:	Stephen Messer
Project Role:	Co-I
Research Identifier (e.g. ORCID ID):	

Nearest person month worked:	2
Contribution to Project	Review of cases and classification of subjects into low and high trauma groups.
Funding Support:	

Name:	Gordon Broderick
Project Role:	Co-I
Research Identifier (e.g. ORCID ID):	
Nearest person month worked:	1
Contribution to Project	Redesign of logic modeling, and oversight and guidance on system level analysis.
Funding Support:	

Name:	Nancy Klimas
Project Role:	Co-I
Research Identifier (e.g. ORCID ID):	
Nearest person month worked:	1
Contribution to Project	Coordination of data transfer.
Funding Support:	

Name:	Ricardo Castellanos
Project Role:	Research Associate/ Grant funded research staff
Research Identifier (e.g. ORCID ID):	
Nearest person month worked:	6
Contribution to Project	Oversees the integration of the data analysis pipeline including univariate and multivariate, and correlation analysis, as well as machine learning and classification schemes.
Funding Support:	

Name:	Rajeev Jaundoo
Project Role:	Research Programmer I/ Grant funded research staff
Research Identifier (e.g. ORCID ID):	
Nearest person month worked:	5
Contribution to Project	Amalgamation of pharmacogenomics information for the identification of drug treatment targets.
Funding Support:	

Name:	Mary Jeffrey
Project Role:	Data Analyst/ Grant funded research staff
Research Identifier (e.g. ORCID ID):	

Nearest person month worked:	6
Contribution to Project	Consolidation of psychological data, review of case assessments, statistical analysis of sub-group profiles.
Funding Support:	

Name:	Ramesh Govindan
Project Role:	Data Analyst/ Grant funded research staff
Research Identifier (e.g. ORCID ID):	
Nearest person month worked:	3
Contribution to Project	Consolidation of biological data.
Funding Support:	

Name:	Rebecca Reinhardt
Project Role:	Research Assistant/ Grant funded research staff
Research Identifier (e.g. ORCID ID):	
Nearest person month worked:	1
Contribution to Project	Assisting in consolidation of psychological data, review of case assessments.
Funding Support:	

Name:	Jessica McDonald
Project Role:	Research Assistant/ Grant funded research staff
Research Identifier (e.g. ORCID ID):	
Nearest person month worked:	1
Contribution to Project	Assisting in consolidation of psychological data, review of case assessments
Funding Support:	

Name:	Zeus Cortes
Project Role:	Volunteer research staff
Research Identifier (e.g. ORCID ID):	
Nearest person month worked:	2
Contribution to Project	Assisting in machine learning and classification schemes.
Funding Support:	Miami-Dade STEM grant

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

- In February 2017 Dr. Broderick moved his laboratory to Rochester General Hosptial in Rochester, NY. At that time DoD-CDMRP award W81XWH-15-1-0582 was

transferred to Dr. Craddock as PI (1.8 calendar) with Dr. Broderick as Co-PI (2.6 calendar), with efforts remaining unchanged from prior to the move.

- Since the initiation of this award Dr. Klimas has had the following changes to her activity:

  Added:
  W81XWH-15-1-0163 (PI Waziry)
  DoD GWIRP
  An Integrated Genomics and Cell Biology Approach to Correlate Novel GWI Indicators of Infections and Neuroinflammatory Mechanisms with Targeted Drug Therapy
  0.216 calendar

  GW160123 (PI Klimas)
  DoD GWIRP
  The Use of B-Cell Depletion Therapy (BCDT) in Gulf War Illness: A Phase 1/2 Study
  0.36 calendar

  GW160051 (PI Salgueiro)
  DoD GWIRP
  Growth Hormone-Releasing Hormone (GHRH) Antagonist: Evaluation of Beneficial Effects for Gulf War Illness
  0.36 calendar

  GW160116 (PI Nathanson)
  DoD GWIRP
  Genomics approach to find female specific mechanisms of GWI pathobiology
  0.216 calendar

  RFA-NS-17-021(PI Klimas)
  NIH
  Deconstructing ME/CFS: Using a computational biology framework to identify ME/CFS subtypes and transform treatment
  1.2 calendar

  Removed:
  R56AI120724 (PI Klimas)
  NIH
  Microbial Discovery and Immunity in ME/CFS
  0.365 calendar

- Since the initiation of this award Dr. Craddock has had the following changes to his activity**:**

  Added:
  GW160116 (PI Nathanson)
  DoD GWIRP
  Genomics approach to find female specific mechanisms of GWI pathobiology
  0.48 calendar

What other organizations were involved as partners?

Organization Name:	Miami-Dade College
Location of Organization:	Miami, Florida, USA
Contribution:	Personnel exchanges: Zeus Cortes is a student in the School of Engineering and Technology at Miami-Dade College. This exchange was made possible by the STEM-EngInE (Science Technology Engineering and Mathematics - Educate, Intern and Employ) program through the Summer Research Institute funded by the Department of Education.

8. SPECIAL REPORTING REQUIREMENTS

Nothing to Report

9. APPENDICES

A. REFERENCES

1. Bourdette DN, et al.(2001) Symptom factor analysis, clinical findings, and functional status in a population-based case control study of Gulf War unexplained illness. *J Occup Environ Med.* 43(12): 1026-40.
2. Craddock, T. J., Fritsch, P., Rice Jr, M. A., Del Rosario, R. M., Miller, D. B., Fletcher, M. A., ... & Broderick, G. (2014). A role for homeostatic drive in the perpetuation of complex chronic illness: Gulf War Illness and chronic fatigue syndrome. *PLoS One*, *9*(1), e84839.
3. Craddock TJ, Harvey JM, Nathanson L, Barnes ZM, Klimas NG, Fletcher MA, et al. Using gene expression signatures to identify novel treatment strategies in gulf war illness. BMC medical genomics. 2015;8(1):36.
4. Do, C. B., & Batzoglou, S. (2008). What is the expectation maximization algorithm? *Nature biotechnology, 26*(8), 897.
5. Fritsch, P., Craddock, T. J., del Rosario, R. M., Rice, M. A., Smylie, A., Folcik, V. A., ... & Broderick, G. (2013). Succumbing to the laws of attraction: Exploring the sometimes pathogenic versatility of discrete immune logic. *Systems Biomedicine*, *1*(3), 179-194.
6. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, *11*(1), 10-18.
7. Kamburov A, Wierling C, Lehrach H, Herwig R. ConsensusPathDB—a database for integrating human functional interaction networks. Nucleic acids research. 2009;37(suppl 1):D623-D8.
8. Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R. ConsensusPathDB: toward a more complete picture of cell biology. Nucleic acids research. 2011;39(suppl 1):D712-D7.
9. Kamburov A, Stelzl U, Lehrach H, Herwig R. The ConsensusPathDB interaction database: 2013 update. Nucleic acids research. 2013;41(D1):D793-D800.
10. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic acids research. 2011:gkr988.
11. Kang HK, et al. (2003) Post-traumatic stress disorder and chronic fatigue syndrome-like illness among Gulf War veterans: a population-based survey of 30,000 veterans. *Am J Epidemiol.* 157(2): 141- 148.
12. McDonald, S. D., Thompson, N. L., Stratton, K. J., & Calhoun, P. S. (2014). Diagnostic accuracy of three scoring methods for the Davidson Trauma Scale among US military Veterans. *Journal of anxiety disorders*, *28*(2), 160-168.
13. Nicolson GL (2001) Gulf War Illnesses: Causes and Treatments. *Armed Force Med Develop.* 2:41-44.
14. Nicolson GL, et al. (2000) Diagnosis and Integrative Treatment of Intracellular Bacterial Infections in Chronic Fatigue and Fibromyalgia Syndromes, Gulf War Illness, Rheumatoid Arthritis, and Other Chronic Illnesses. *Clin Pract Altern Med.* 1:92-106.
15. Pacella ML, et al. (2013) The physical health consequences of PTSD and PTSD symptoms: a meta-analytic review. *J Anxiety Dis.* 27(1): 33-46.

16. Suthram S, Dudley JT, Chiang AP, Chen R, Hastie TJ, Butte AJ. Network-based elucidation of human disease similarities reveals common functional modules enriched for pluripotent drug targets. PLoS Comput Biol. 2010;6(2):e1000662.
17. Unwin C, et al. (1999) Health of UK servicemen who served in Persian Gulf War. *Lancet*. 353:169-78.
18. Weiner MW, et al. (2011) The relationship between Gulf War illness, brain N-acetylaspartate, and post-traumatic stress disorder. *Mil Med.* 176(8):896-902.

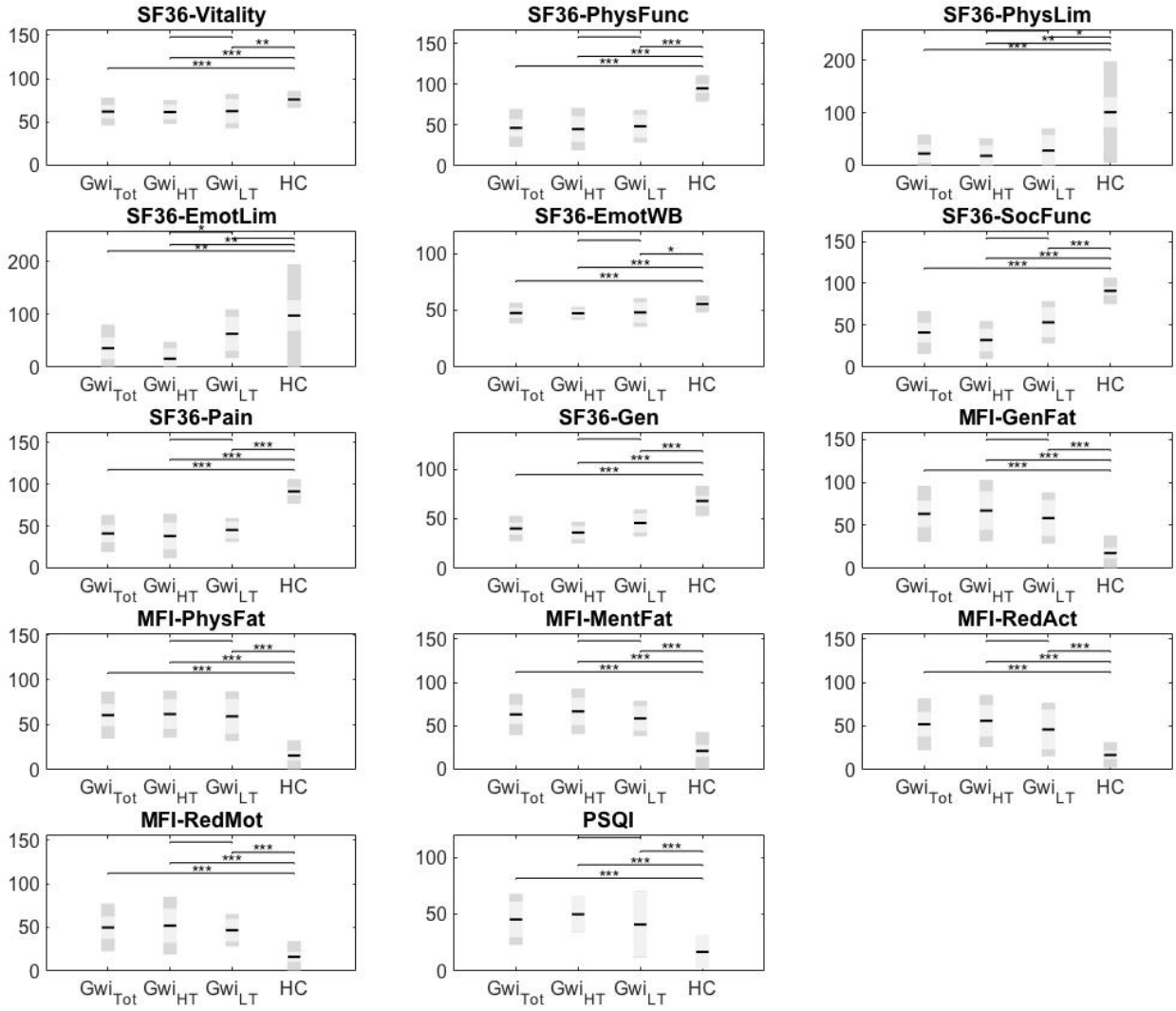

Figure 1: Two sample t-test comparisons between female healthy controls (HC; n=46), female GWI in total (GWI_{Tot}; n=19), and GWI subgroups based on presence (GWI_{HT}; n=11) and absence (GWI_{LT}; n = 8) of trauma. * p<=0.05; ** p<=0.01; *** p<=0.001.

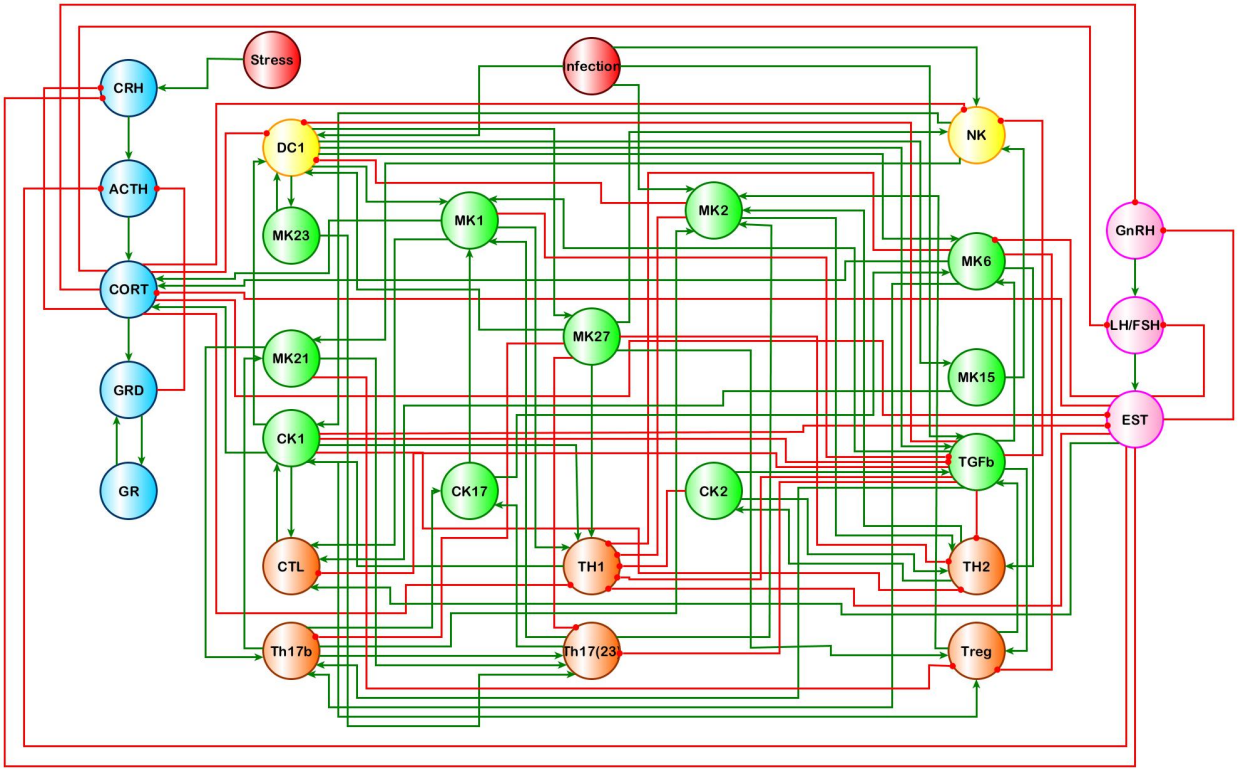

Figure 2. Aggregate multi-axis model with negative estrogen feedback to HPA and HPG. Integration into a comprehensive model of the coarse HPA-HPG-immune circuitry for females (Craddock et al., 2014) with our fine-grained circuit model of immune signaling Fritsch et al., 2013). This aggregate model is now being used as a basis for continued development. As before, green directed edges represent an up-regulation of the target by the source node whereas a red terminal edge represents a suppressive action.

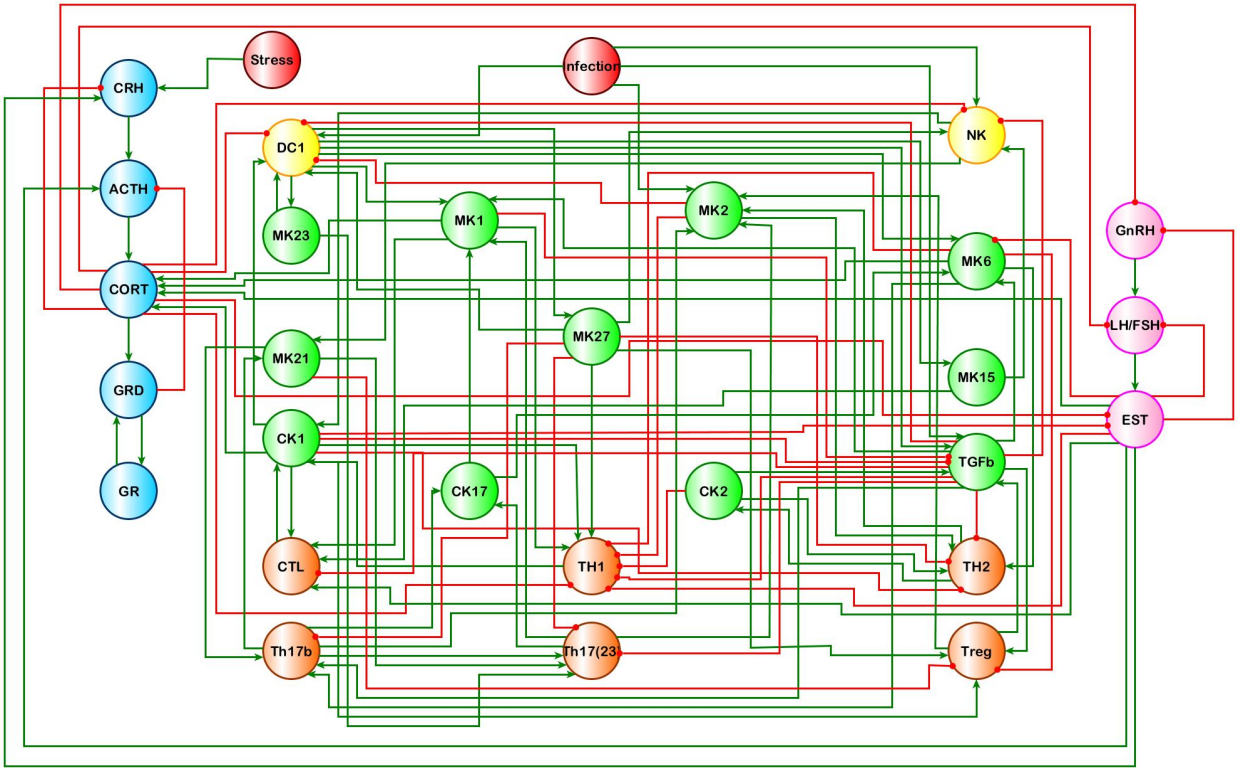

Figure 3. Aggregate multi-axis model with positive estrogen feedback to HPA and negative estrogen feedback to HPG. Integration into a comprehensive model of the coarse HPA-HPG-immune circuitry for females (Craddock et al., 2014) with our fine-grained circuit model of immune signaling Fritsch et al., 2013). This aggregate model is now being used as a basis for continued development. As before, green directed edges represent an up-regulation of the target by the source node whereas a red terminal edge represents a suppressive action.

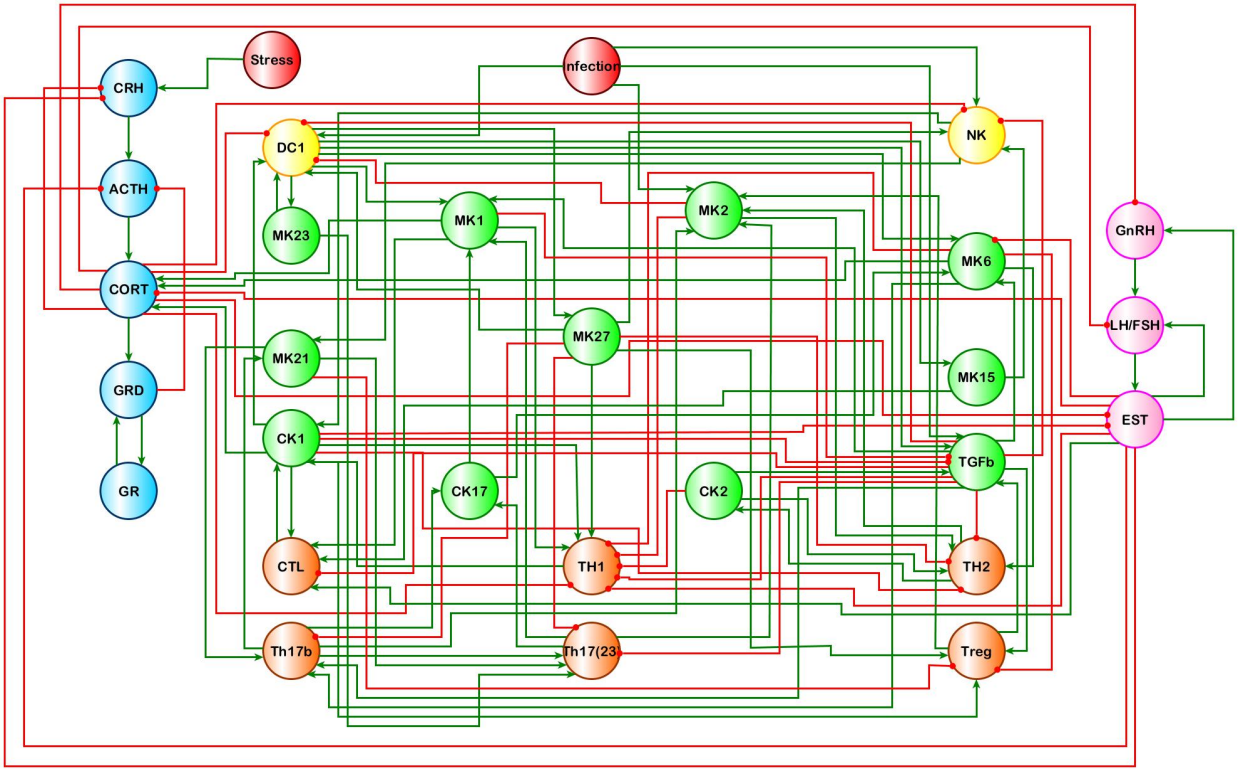

Figure 4. Aggregate multi-axis model with negative estrogen feedback to HPA and positive estrogen feedback to HPG. Integration into a comprehensive model of the coarse HPA-HPG-immune circuitry for females (Craddock et al., 2014) with our fine-grained circuit model of immune signaling Fritsch et al., 2013). This aggregate model is now being used as a basis for continued development. As before, green directed edges represent an up-regulation of the target by the source node whereas a red terminal edge represents a suppressive action.

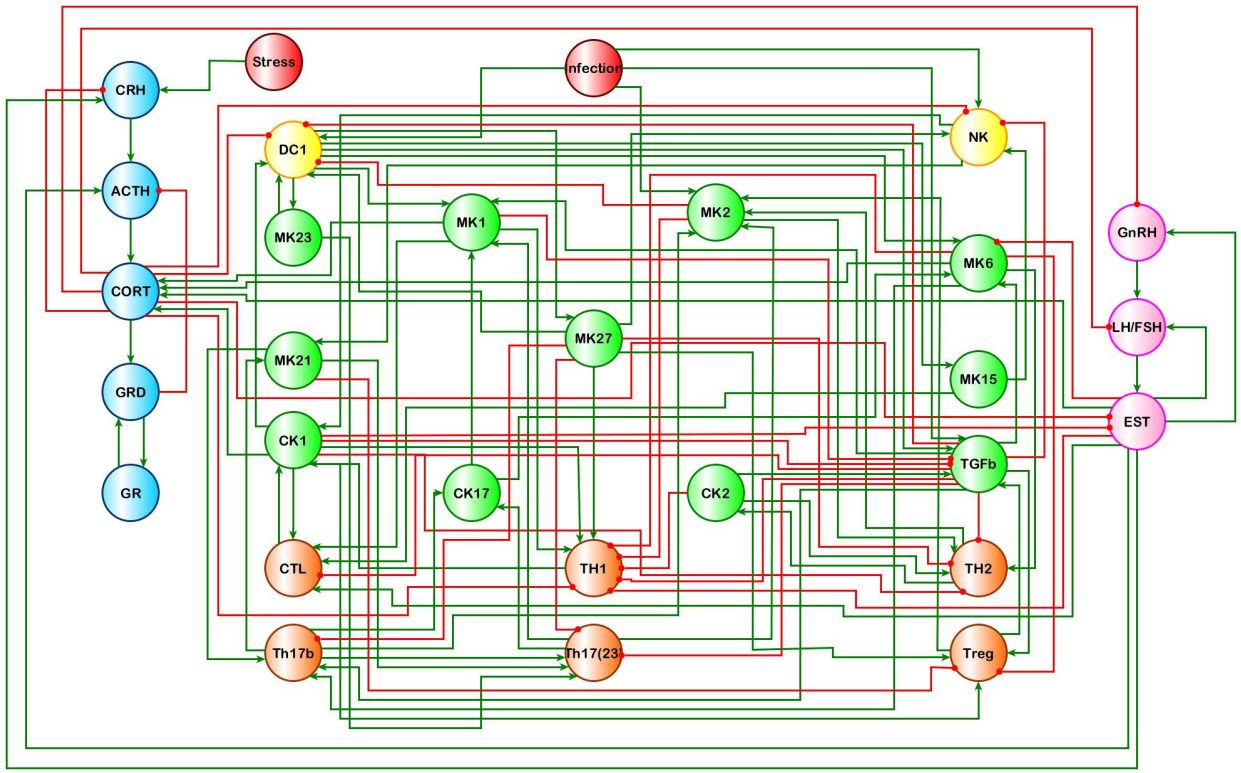

Figure 3. Aggregate multi-axis model with positive estrogen feedback to HPA and HPG. Integration into a comprehensive model of the coarse HPA-HPG-immune circuitry for females (Craddock et al., 2014) with our fine-grained circuit model of immune signaling Fritsch et al., 2013). This aggregate model is now being used as a basis for continued development. As before, green directed edges represent an up-regulation of the target by the source node whereas a red terminal edge represents a suppressive action.

Application of Machine Learning Techniques to Classify High/Low Trauma Gulf War Illness Patients Against Healthy Controls

Zeus Cortes-Aguila[1], Ricardo Castellanos[2,3], Rajeev Jaundoo[2,3], Travis J. A. Craddock[2,3,4,5]

[1]School of Engineering and Technology, Miami-Dade College, Miami, FL
[2]Institute for Neuro-Immune Medicine, Nova Southeastern University, Ft. Lauderdale, FL
[3]Department of Psychology and Neuroscience, Nova Southeastern University, Ft. Lauderdale, FL
[4]Department of Computer Science, Nova Southeastern University, Ft. Lauderdale, FL
[5]Department of Clinical Immunology, Nova Southeastern University, Ft. Lauderdale, FL

1 Introduction

Gulf War Illness (GWI) is a chronic multi-symptom disorder which affects veterans of the 1991 Persian Gulf War [Craddock et al. 2015]. This disorder includes chronic symptoms such as headaches, fatigue, musculoskeletal pain, and respiratory and cognitive problems among others [Craddock et al. 2015]. There is currently no known cure available for GWI, although several studies suggest that enhancing neurogenesis and suppression of the inflammatory response alleviates the mood and cognitive dysfunction observed in GWI[2]. Since there are no widely accepted biomarkers for GWI, veterans are primary diagnosed based on psychological or psychiatric evaluations. Furthermore, many veterans that suffer with GWI are often also diagnosed with post-traumatic stress disorder (PTSD) [citation][3]. This adds an additional layer of complexity for GWI diagnosis due to its strong overlap with PTSD, which may affect potential treatment avenues.

Machine learning techniques are used to identify complex patterns and perform classification of data [citation]. In our study, we classified GWI and healthy control data, including some variations with trauma scores. The decision tree algorithm J48 was used on the patient data due to its ability to output both confusion matrices and visualized decision trees. These decision trees enabled us to distinguish which cytokines had the greatest impact on machine learning classifications. Subjects with PTSD exhibit immunological dysfunction, including increased levels of circulating cytokines (eg. IL-1B, IL-2, IL-4, IL-6, IL-8, IL-10 and TNF-a), suppressed natural killer cell activity, and lower total T lymphocyte counts when compared to healthy controls. As GWI has also been shown to have similar immunological underpinnings, this suggests that a screen for probable PTSD diagnosis may delineate at least two phenotypes of GWI. The objective is to gather biological data of the circulating cytokines for patients diagnosed with GWI and healthy controls, then run machine learning algorithms to detect underlying correlations within the data.

2 Methods & Procedures

We obtained our data from the Institute for Neuro-Immune Medicine (INIM) where patients were tested at three different times: at rest (T0), at peak exercise (T1), and during post-exercise recovery (T2). The cytokines measured during these times included NPY, CORT, IL-1a, IL-1b, IL-2, IL-4, Il-5, IL-6, IL-8, IL-10, IL-12, IL-13, IL-15, IL-17, IL-23, IFNg, TNFa, and TNFb. Additionally, we defined a trauma score for each patient, where any score equal to or above 70 indicates high trauma, and scores below 70 are considered low trauma. We used the machine learning program WEKA 3.8 to perform our analysis[4].

We first separated the data into several different groups to determine which group yielded the best possible correct classification percent. The groups were prepared as following: GWI vs healthy control (HC), high trauma GWI (HTG) vs HC, low trauma GWI (LTG) vs HC, and finally HTG vs LTG vs HC. All groups were tested during each recorded time (T0, T1, T2) as well as during all times together. Next, Weka was used to perform unbiased supervised machine learning to predict the classification of our chosen group. Here, the unpruned J48 decision tree algorithm with 10-fold cross validation was employed. Cross-validation is a technique to evaluate predictive models by randomly partitioning the original sample into multiple training set and test set[5]. The training set is used to "teach" the model while the test set validates the results. Once the cross validation was complete, we extracted information such as percentage of correctly classified instances, confusion matrices, and the visual trees.

4 Results

With the given data for a majority of the different combinations of datasets, Weka could predict whether or not a patient suffers GWI better than 50%. The best overall criteria for prediction was in Low Trauma GWI vs the healthy control post exercise challenge (T2).

	T0	T1	T2	All Times
GWI vs HC	59.25%	51.85%	57.41%	62.96%
HTG vs HC	61.90%	47.62%	61.90%	59.52%
LTG vs HC	69.23%	58.97%	74.36%	56.41%
HTG vs LTG vs HC	44.44%	31.48%	44.44%	46.30%

Table 1: GWI Correctly Classified Instance Percentages at varying times.

Another important result gathered from the visualization tree was the cytokines that were the deciding factor in classifying the data. By adding up all of the deciding classifications each cytokine did and dividing by the total number, we arrived at a percentage of the total classified shown in Figure 2.

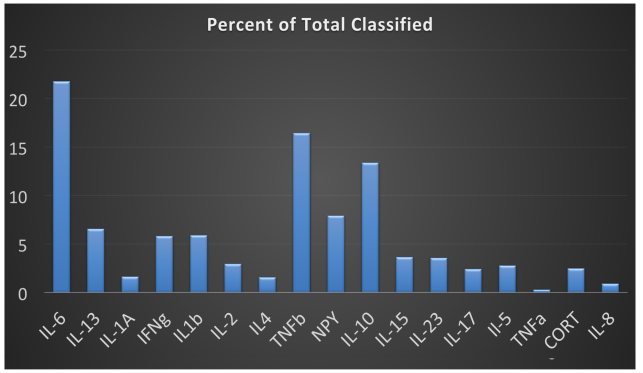

Figure 2: Percentage of Total Classified by each cytokine.

5 Discussion

The results indicate that under the certain circumstances it is possible to use machine learning techniques to arrive at an accuracy of prediction 24.36% higher than chance. The data gathered from the visual trees also gives insight on the potential biomarkers associated with GWI. It is worth noting that Il-6, Il-10, and TNFb all have some connection to inflammatory response. These results of inflammatory-related cytokines are also backed by another study that indicates at inflammation is a component of the pathobiology of GWI[6]. Furthermore, the results indicate that there is a difference between low trauma and high trauma groups when it comes to classifying against healthy controls.

6 Conclusions

The results of this experiment show that by utilizing trauma as a means to categorize the patient, the accuracy of classification increases. It should be mentioned that our initial dataset consisted of 151 patients. Using machine learning algorithms with a larger dataset will increase the confidence level of our results and possibly increase accuracy, so further trials of this experiment may be repeated when more patient data is collected.

References

1. Craddock et al. PLoS ONE 7(6): e37251. (2012)
2. Parihar V. K., et al. *Neuropsychopharmacology* 38 2348–2362. 10.1038/npp.2013.158 (2013)
3. Nicolson et al. *Clin Pract Altern Med.* 1:92-106. (2000)
4. Eibe Frank, et al. (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.
5. Vanschoren et al . SIGKDD Explorations 15(2), pp 49-60, (2013)
6. Johnson GJ et al. . PLoS ONE 11(6): e0157855. (2016)

Acknowledgements

The original research summarized in the following was supported, in part, by U.S. Department of Education grant award P031C110011, and awards by the U.S. Department of Defense (W81XWH-16-1-0632 and W81XWH-16-1-0552 (Craddock PI)). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the respective funding agency.

Contact Information

Contact: zeuscortes97@gmail.com
tcraddock@nova.edu
URL: www.nova.edu/inim-csb